

# Reporte de Análisis: Correlaciones No Lineales

## Introducción

El presente trabajo tuvo como objetivo evaluar la relación entre distintas variables clave de propiedades listadas en Airbnb mediante el uso de modelos de regresión no lineal. Se analizaron seis pares de variables utilizando dos modelos matemáticos distintos: uno logarítmico y otro exponencial. Los resultados obtenidos permiten comprender mejor cómo interactúan estas variables en el contexto del mercado de alquileres temporarios.

## Metodología

Se aplicaron dos modelos matemáticos a cada par de variables:

1. **Modelo Logarítmico:**  $y = a \cdot \ln(x) + b$
2. **Modelo Exponencial:**  $y = a \cdot e^{(-bx)} + c$

La efectividad de cada modelo se midió mediante el coeficiente de determinación ( $R^2$ ) y su correspondiente coeficiente de correlación ( $r = \sqrt{R^2}$ ). Se analizaron los siguientes pares de variables:

1. Tasa de respuesta del anfitrión vs Número de reseñas
2. Tasa de aceptación del anfitrión vs Número de reseñas
3. Cantidad de propiedades del anfitrión vs Número de reseñas
4. Capacidad de alojamiento vs Número de habitaciones
5. Reseñas por mes vs Puntuación de reseñas
6. Precio vs Número de camas

El análisis más significativo fue el que relacionó la capacidad de alojamiento (accommodates) con el número de habitaciones (bedrooms). El modelo exponencial mostró un  $R^2$  de 0.324, lo que indica que explica aproximadamente el 32% de la variabilidad en los datos. La correlación resultante ( $r = 0.569$ ) sugiere una relación moderada entre estas variables.

Este resultado sugiere que, si bien existe una relación positiva entre el número de habitaciones y la capacidad de alojamiento, esta no es estrictamente lineal. Es probable que propiedades con más habitaciones tiendan a tener una mayor capacidad, pero con rendimientos decrecientes.

Los modelos aplicados a las siguientes relaciones mostraron resultados más débiles. Para la tasa de respuesta (host\_response\_rate), ambos modelos arrojaron

un  $R^2$  de aproximadamente 0.036, indicando una relación muy débil. En el caso de la tasa de aceptación (`host_acceptance_rate`), el modelo exponencial mostró un ligero mejor desempeño ( $R^2 = 0.063$ ) que el logarítmico.

Estos resultados sugieren que el número de reseñas tiene una influencia limitada en el comportamiento de respuesta y aceptación de los anfitriones. Factores no considerados en este análisis podrían ser más determinantes en estas tasas.

Los análisis restantes mostraron relaciones particularmente débiles o prácticamente nulas:

- La cantidad de propiedades de un anfitrión no mostró relación significativa con su número de reseñas ( $R^2 < 0.013$ )
- La frecuencia de reseñas mensuales no se correlacionó con las puntuaciones recibidas ( $R^2 < 0.006$ )
- El precio no mostró relación con el número de camas ( $R^2 < 0.002$ )

Los resultados indican que estas variables probablemente están influenciadas por otros factores no considerados en el análisis actual, o que sus relaciones no siguen patrones logarítmicos o exponenciales simples.

## Conclusiones

En conclusión, el modelo exponencial demostró ser más efectivo que el logarítmico en este análisis, mostrando un mejor desempeño en cuatro de los seis casos estudiados y destacándose particularmente en la relación entre capacidad de alojamiento y número de habitaciones, donde alcanzó un  $R^2$  de 0.324 frente al 0.015 del modelo logarítmico. Aunque ambos modelos presentaron resultados similares en la mayoría de los análisis, el exponencial mostró mayor robustez al manejar valores cercanos a cero y una ligera ventaja al predecir la tasa de aceptación de reservas. Sin embargo, los bajos valores de  $R^2$  obtenidos en cinco de los seis casos ( $\leq 0.1$ ) indican que ninguno de los dos modelos logra explicar satisfactoriamente estas relaciones por sí solo, sugiriendo la necesidad de explorar modelos alternativos más complejos o incluir variables adicionales que puedan capturar mejor la dinámica subyacente en los datos de Airbnb. Este estudio resalta la importancia de evaluar múltiples enfoques metodológicos y contextualizar los resultados, recomendándose el modelo exponencial como punto de partida para futuros análisis, pero considerando su complementación con otras técnicas cuando las correlaciones sean débiles.